学校编码：10384　　　　分类号______密级______
学号：19020071152099　　　　UDC______

厦门大学

硕 士 学 位 论 文

VaR计算方法的研究

The research about the calculation of Value at Risk

吴 明 州

指导教师姓名：　刘 继 春 教 授
专 业 名 称：　概 率 论 与 数 理 统 计

论文提交日期：　2010 年　 月
论文答辩时间：　2010 年　 月
学位授予日期：　2010 年　 月

答辩委员会主席：__________
评　　阅　　人：__________

2010 年　 月

厦门大学学位论文原创性声明

兹呈交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明。本人依法享有和承担由此论文而产生的权利和责任。

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　　）1.经厦门大学保密委员会审查核定的保密学位论文，于　　年　　月　　日解密，解密后适用上述授权。

（　　　）2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

摘要

VaR(Value at Risk)是度量、识别和控制金融市场风险的重要工具. 本文我们在讨论了基于GARCH$(1,1)$模型的VaR和极值理论VaR的计算后, 假设资产对数收益率模型为:

$$r_u = \theta_u S_{u-1}^{\beta_u} \varepsilon_u + \tau_u$$

针对其中的误差项ε_u的分布, 我们提出用正态逆高斯分布(Normal inverse gaussian NIG)和双曲分布(Hyperbolic distribution HYP)代替常用的正态分布或t分布, 建立半参数随机波动模型. 利用极大似然方法, 估计出参数θ_u, β_u和τ_u. 从而得到波动σ_u的估计, 计算在给定0.05和0.01概率条件下当ε_u服从均值为0, 方差为1的NIG分布和HYP分布时对应的VaR值.

关键词：NIG, HYP, VaR, 半参数.

Abstract

VaR(Value at Risk) is a fundamental tool for measuring ,recognizing and controlling the risk of financial markets. In this dissertation,first we discuss the VaR calculation base on the GARCH$(1, 1)$ model under different residuals distribution,then the VaR calculation base on the Extreme Values Theory. After this we make a hypothesis that the logarithmic return model is

$$r_u = \theta_u S_{u-1}^{\beta_u} \varepsilon_u + \tau_u$$

and we suggest that the residuals of this model is follow Normal inverse gaussian (NIG) or Hyperbolic distribution (HYP) instead of Normal distribution and Student't distribution. In the end,we use the semiparametric approach and pseudo-likelihood methods to estimate the volality,calculate the VaR in probability of 0. 01 and 0. 05 which under the hypothesis that the distribution of residuals are Normal inverse gaussian (NIG) or Hyperbolic distribution (HYP) with mean 0 and variance 1

Key Words: NIG, HYP, VaR, Semiparametric.

目　录

Contents

表格

插图

第一章　引言

1.1 研究背景与VaR研究方法现状

自20世纪70年代以来, 由于金融自由化, 信息技术与金融创新和金融管制的放松等因素的影响, 金融市场的波动性明显增强, 金融体系的稳定性不断下降, 金融市场的风险与日俱增. 金融风险不仅严重影响了金融机构和工商企业的正常营运, 而且还对一国乃至全球金融及经济的稳定构成了严重威胁, 频繁发生的金融危机更是带来了一系列的严重后果. 1994年的墨西哥比索的贬值, 引发一场波动全球的金融危机; 1997年的亚洲金融危机给东南亚各国的经济带来了沉重的打击. 世界许多金融机构由于对市场风险管理不善而导致巨大损失. 因此, 金融风险引起了全世界的金融界、企业界、政府当局和国际金融组织的密切关注和高度重视. 金融机构这些危机也使监管当局频频出台新的政策, 巴塞尔银行监管委员会于2004年6月推出的《巴塞尔新资本协议》, 更是对金融机构的风险管理提出了严格的要求.

金融市场风险管理, 包括风险识别、风险度量、风险管理决策与实施, 以及风险控制四个阶段. 其中风险度量是金融风险管理过程中最重要的一个环节, 它包括衡量各种风险导致损失可能性的大小以及损失发生的范围和程度. 度量金融市场风险提出过许多方法, 其中风险价值(Value-at-Risk, 缩写VaR)方法由于其直观性、简洁性和科学性, 因此是很受欢迎的方法, 它能将不同交易、不同业务部门的市场风险集成为一个数值, 因而易于理解和管理. VaR方法不仅能够清晰地描述市场风险的大小, 而且还有严谨的概率统计理论作依托, 因而基于VaR的金融风险度量与管理方法得到了国际金融界的广泛支持和认可, 并使得机构投资者和市场监管者能够很方便地运用它进行金融市场风险判断. 目前, VaR已发展成为现代风险量化管理中应用最广泛、最重要的方法, 它们在证券投资组合的风险测量、经营机构(证券经营机构, 其它金融机构甚至一般的公司)的整体风险监控中得到越来越广泛的应用, 关于VaR更加详细的介绍可以参阅Duffie, Pan(1997).

计算VaR首先必须了解金融数据的特征, 只有正确认识和了解金融市场数据的特征事实, 才能正确地建立模型, 运用合适的模型度量VaR对实际才有指导意义. 金融市场的数据具有集聚性, 尖峰厚尾, 不对称等特性.通过研究, 人们发现金融时间序列的波动呈现出时变性, 即波动不是固定不变的, 具有异方差的特点. 所有传统的VaR模型都是部分围绕或全部围绕这些经验规律来建立模型的. 研究者们从理论和应用方面做了许多工作. 理论方面人们研究VaR的方法有很多. Manganelli, Engle(2001)对这些

方法做了总结比较, 现在的VaR计算主要分为三大类:参数型如So, Yu(2006), 非参数型如Cai, Wang(2008)和Chen, Hardle, 和Jeong(2008), 半参数型如Fan, Yao(2003).

1.2 本文的结构和内容

VaR的计算涉及到未来市场分布的选择, 波动性估计和资产定价模型等三个方面. 本学位论文包括三部分:基于GARCH$(1,1)$随机波动模型的VaR, 极值VaR的计算和半参数随机波动模型VaR.

在第二章, 我们将讨论GARCH$(1,1)$模型下, 当ε_t在正态分布, t分布, 偏t分布(Skew Student's t distribution SKST)[22]和偏广义误差分布(Skew generalized error distribution SGED)[29]的假设条件下的各种VaR值的计算. 通过研究, 人们发现金融时间序列的波动呈现出时变性, 即波动不是固定不变的, 具有异方差的特点.正是由于这些特点Engle提出ARCH模型分析时间序列的异方差性后, Bollerslev(1986)提出了GARCH模型, GARCH$(1,1)$ 模型在实证研究中已取得巨大成功, 被许多计量经济学家认为是基准模型. 对于GARCH模型中的ε_t一般假设服从正态分布, Hansen(1994)最早提出了偏t分布(SKST), 并用它来对收益率进行建模. 与t分布相似SKST分布有着较厚的尾部, 但它弥补了t分布是对称分布的不足, 从而在一定程度上提高波动估计的效率. Nelson,Cao(1992)建议在金融时间序列建模中使用广义误差分布(GED), Lehnert(2003)提出了一种更一般的分布, 带偏的广义误差分布(SGED), 这也是对GARCH模型中ε_t分布假设的一种改进. 本章我将分别在ε_t分布服从正态分布, t分布, 偏t分布(SKST)和偏广义误差分布(SGED)的假设条件下, 以沪深300指数日对数收益率为数据, 建立波动估计模型, 从而计算出不同分布下的VaR值.

在第三章中, 我们将讨论的是统计学中怎样用极值理论来描述一些金融市场出现的极端交易情况, 以及如何利用极值理论在计算VaR. Tsay(2005)对极值理论计算VaR做了详细的叙述. 极大的价格波动在金融市场是比较少见的的, 但每次出现其影响是巨大的, 他有可能造成金融机构事先由于未做好从分准备, 从而导致破产. 如2007年5月印花税的突然调整对A股市场造成的巨大冲击, 2009年7月底八月初A股市场那波调整, 在那个阶段市场的不理智造成了股指的巨大波动. 我们仍以沪深300指数日对数收益率为数据源, 首先利用一般广义极值分布(GEV)计算沪深300指数日VaR. 其次, 采用高峰限值法(Peaks-over-Threshod缩写POT), 该方法首先给出一个阈值, 将超出已选定的阈值的值选定为极值, 利用帕累托分布计算沪深300指数日VaR.

在统计学中, 一个模型$\{P_\theta, \theta \in \Theta\}$称作参数模型是指这个模型的索引参数$\theta$是一个有限维向量, 即$\theta$取值所成的集合是$R^k$的一个子集, 而当$\theta$的取值所成集合是某些空

间的子集, 并不要求是有限维的, 例如我们可以考虑那些所有均值为零的那些分布构成的集合, 这样的空间是一个带有拓扑结构的向量空间, 但也许不是一个有限维的空间. 因此, $\Theta \in \mathscr{F}$, 这里$\mathscr{F}$是某个维数无线的空间, 而当参数空间$\Theta \subset R^k \times \mathscr{F}$, 即参数包含了有限维的部分和无限维的部分, 我们称这样的模型为半参数模型.

在第四章中, 我们将假设沪深300日对数收益率满足模型$r_u = \theta_u S_{u-1}^{\beta_u}\varepsilon_u + \tau_u$, 其中$\varepsilon_t$分别服从均值为0方差为1的NIG分布和HYP分布, 从而对随机波动σ_t 建立半参数模型, 得到$\widehat{\sigma_t}$作为σ_t的估计, 然后计算在分别给定概率时对应的VaR值. 并与第二章和三章模型得到的结果进行比较.正态分布是建立我们在建模中最常用的分布, 他有着密度函数简单, 参数少(只有μ, σ)和计算简单等优点, 但由于金融数据本身尖峰厚尾的特性, 人们发现采用正态分布来对金融数据进行建模产生了很多明显的不足.为此有不少研究者对正态分布假设提出改进, 如LAPLACE分布, CAUCHEY分布, 混合正态分布等等, 这些研究工作都在一定程度上提高了模型的精度, 但由于上述分布有可能只能对某些特定的金融数据适用, 当数据类型发生改变时, 就有可能凸显出它们的不足.Barndorff-Nielsen(1977)提出了一种非常一般的分布族, 广义双曲分布(General hyperbolic distribution GHD). 这是一类非常一般的分布, 其密度函数有$(\lambda, \alpha, \beta, \delta, \mu)$五个参数.许多参见的分布如正态分布, t分布都是它的特殊情况.在GHD分布中应用最广的是正态逆高斯分布(Normal inverse gaussian NIG)和双曲分布(Hyperbolic distribution HYP), 这两个分布在描述金融数据时有着十分良好的变现.

第二章　基于GARCH$(1,1)$随机波动模型的VaR

GARCH$(1,1)$随机波动模型的误差项ε_t在正态分布, t分布和GED分布假设下的VaR模型在实证中不断地暴露其不足之处, 它们不能很好地捕捉到金融数据的偏度, 尖峰和厚尾的特征. 在本章中, 我们将以沪深300指数日对数收益率为数据, 计算当误差项ε_t服从SKST分布和SGED分布的VaR, 并与正态分布, t分布假设下的VaR结果进行比较.

2.1 基本定义

在Engle提出ARCH模型分析时间序列的异方差性后, Bollerslev(1986)又提出了GARCH模型. GARCH模型是一个专门针对金融数据的回归模型, 他对误差的方差进行了进一步建模, 特别适用于波动性的分析和预测.

定义2.1.1[38] 具有$p \geq 1$和$q \geq 0$的广义自回归条件异方差模型为：

$$r_t = \mu + h_t \tag{2.1.1}$$

$$h_t = \sigma_t \varepsilon_t\ ,\ \varepsilon_t | \Omega_{t-1} \sim IID(0,1) \tag{2.1.2}$$

$$\sigma_t^2 = C_0 + \sum_{i=1}^{p} b_i h_{t-i}^2 + \sum_{i=1}^{q} a_j \sigma_{t-j}^2 \tag{2.1.3}$$

其中$C_0 \geq 0$, $b_i \geq 0$和$a_j \geq 0$是常数, $\{\varepsilon_t\} \sim IID(0,1)$, 且$\forall t$, ε_t与$\{X_{t-k}, k \geq 1\}$相互独立, 则称随机过程$\{r_t\}$为GARCH(p,q)过程. 对$p=1$和$q=1$时的GARCH$(1,1)$模型在实证研究中已取得巨大成功, 被许多计量经济学家认为是基准模型.

上述模型中r_t指的是某资产在t时刻的收益率, 在这里我们给出r_t的确切定义.

定义2.1.2 令t时刻的资产价格为S_t, t时刻的资产收益率为r_t, 其中$r_t = log(\frac{S_t}{S_{t-1}})$, 未来$t+1$时刻的资产最低收益率为$r_{t+1}^*$.

定义2.1.3[13] 在一般市场条件下(多头市场)给定置信水平α, 在给定持有期$\mathscr{L}$内, 某一投资组合预期可能发生的最大损失表示为：

$$P\{\triangle V(\mathscr{L}) \leq VaR\} = \alpha, \tag{2.1.4}$$

这里$\triangle V(\mathscr{L})$表示在t到$t+1$时刻资产的价值损失, $\triangle V(\mathscr{L}) \leq 0$, 称VaR为置信水平$\alpha$下处于风险中的价值(Value at Risk).

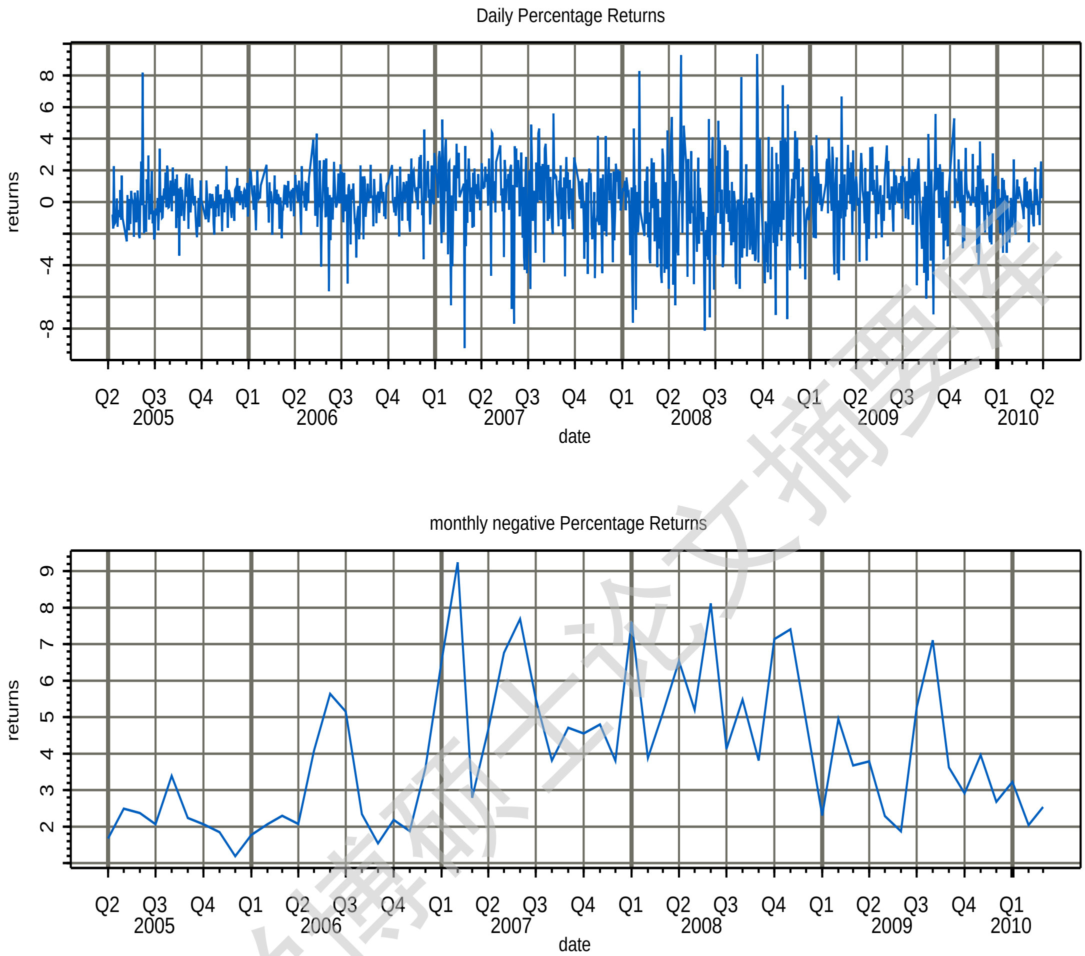

图 2.1 沪深300日收益率与沪深300月最大负收益率

沪深300指数是由上海和深圳证券市场中选取300只A股作为样本编制而成的成份股指数, 它样本覆盖了沪深市场六成左右的市值, 具有良好的市场代表性.沪深300指数是沪深证券交易所第一次联合发布的反映A股市场整体走势的指数. 它的推出, 丰富了市场现有的指数体系, 增加了一项用于观察市场走势的指标, 有利于投资者全面把握市场运行状况, 也进一步为指数投资产品的创新和发展提供了基础条件.

2.2 SKST分布和SGED分布VaR计算

我们考虑如下GARCH$(1,1)$模型, r^*_{t+1} 为未来可能最大损失, r_t为随机波动, $\{\varepsilon_t\} \sim IID(0,1)$.

$$r_t = \mu + h_t, \qquad h_t = \sigma_t \varepsilon_t,\ \varepsilon_t|\Omega_{t-1} \sim IID(0,1), \tag{2.2.1}$$

$$\sigma_t^2 = C_0 + b_0 h_{t-1}^2 + a_0 \sigma_{t-1}^2. \tag{2.2.2}$$

在给定置信水平α时, 有:

$$P(r_{t+1} \le r^*_{t+1}) = P(\frac{r_{t+1}-\mu}{\sigma_{t+1}} \le \frac{r^*_{t+1}-\mu}{\sigma_{t+1}}) = \alpha \tag{2.2.3}$$

则

$$r^*_{t+1} = q(\alpha)\sigma_{t+1} + \mu, \tag{2.2.4}$$

其中$q(\alpha)$是ε_t的α下侧分位数. 因此, 由上述的VaR的定义,我们知$VaR = S_t(\exp\{r^*_{t+1}\} - 1)$.
为了叙述方便, 我们下面给出SKST分布和SGED分布的定义.

定义2.2.1[22] 密度函数为:

$$f(x \mid \lambda, \nu) = \begin{cases} bc(1 + \frac{1}{\nu-2}(\frac{bx+a}{1-\lambda})^2)^{-\frac{\nu+1}{2}}, & x < -\frac{a}{b} \\ bc(1 + \frac{1}{\nu-2}(\frac{bx+a}{1+\lambda})^2)^{-\frac{\nu+1}{2}}, & x \ge -\frac{a}{b}, \end{cases} \tag{2.2.5}$$

分布称为偏t分布(SKST), 其中$2 < \nu < \infty$, $-1 < \lambda < 1$, 且满足$a \equiv 4\lambda c \frac{\nu-2}{\nu-1}$, $b \equiv 1 + 3\lambda^2 - a^2$, $c \equiv \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)\Gamma(\frac{\nu}{2})}}$. 这是一个服从SKST分布的均值为零, 方差为1的随机变量的密度函数. 当$\nu = 0$时, SKST分布就是一般t分布. 更多关于SKST分布的性质可以参看文献Fernandez, Mark.F(1998).

定义2.2.2[29] 密度函数为:

$$f(x \mid \lambda, \nu) = C\exp(-\frac{|x-\delta|^\nu}{[1 - \text{sign}(x-\delta)\lambda]^\nu \theta^\nu}) \tag{2.2.6}$$

分布称为偏广义误差分布(SGED), 其中$C = \frac{\nu}{2\theta\Gamma(\frac{1}{\nu})}$, $\theta = \Gamma(\frac{1}{\nu})^{\frac{1}{2}}\Gamma(\frac{3}{\nu})^{\frac{1}{3}}S(\lambda)^{-1}$, $\delta = 2\lambda A S(\lambda)^{-1}$, $S(\lambda) = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}$, 且$A = \Gamma(\frac{2}{\nu})\Gamma(\frac{1}{\nu})^{-\frac{1}{2}}\Gamma(\frac{3}{\nu})^{-\frac{1}{3}}$. 更多关于SGED分布的性质可以参看文献Lehnert(2003).

命题2.2.3 金融市场上$t+1$时刻的波动σ_{t+1}, 在给定Ω_t条件下有

$$E(\sigma_{t+1}^2 \mid \Omega_t) = C_0 + (a_0+b_0)\sigma_t^2.$$

证明：由于

$$h_t = \sigma_t \varepsilon_t \ , \ \varepsilon_t | \Omega_{t-1} \sim IID(0,1).$$

因此,

$$\begin{aligned} E(\sigma_{t+1}^2 \mid \Omega_t) &= E(C_0 + b_0 h_t^2 + a_0 \sigma_t^2 \mid \Omega_t) \\ &= E(C_0 + b_0 \sigma_t^2 \varepsilon_t^2 + a_0 \sigma_t^2 \mid \Omega_t) \\ &= C_0 + (b_0 + a_0) E(\sigma_t^2 \mid \Omega_t) \\ &= C_0 + (b_0 + a_0)\sigma_t^2. \end{aligned}$$

这样$t+1$时刻的波动$\widehat{\sigma_{t+1}} = \widehat{C_0} + (\widehat{b_0} + \widehat{a_0})\sigma_t^2$.

在假设$\varepsilon_t|\Omega_{t-1} \sim IID(0,1)$下, 沪深300VaR计算的步骤如下：

(1)选取沪深300指数日对数收益率, 令$r_t =$沪深300日对数收益率$\times 100$.

(2) 假设$\varepsilon_t|\Omega_{t-1}$服从均值为0, 方差为1的SKST分布或者服从均值为0, 方差为1的SGED分布. 首先建立GARCH$(1,1)$模型, 然后利用极大似然方法, 估计出参数μ, C_0, b_0, a_0, 得到$t+1$时刻的波动σ_{t+1}.

(3)给定不同的置信水平α.

(4)计算$VaR = S_t(\exp\{\frac{r_{t+1}^*}{100}\} - 1)$, 并画出残差QQ图和在各种不同分布假设下0.05和0.01概率下VaR估计与真实沪深300日最大跌幅对比图.

由表2-1和图2.3至图2.6我们可以发现, 当$\alpha = 0.05$时, 四种不同分布的VaR值无太大差异, 而当$\alpha = 0.01$时, SGED分布和SKST分布的VaR值较正态分布和t分布有了比较大的改进, 且我们通过拟合残差的QQ图2.2, 基于SGED分布和SKST分布的GARCH$(1,1)$模型的效果要明显优于其他两种分布.